

다속성 태도 모델과 협업적 필터링 기반 장소 추천 연구

(A Study on Recommendation Systems based on User multi-attribute attitude models and Collaborative filtering Algorithm)

안병익*, 정구임**, 최혜림***

(Byung-Ik Ahn, Ku-Imm Jung, Hae-Lim Choi)

요약

스마트폰이나 태블릿 PC와 같이 GPS를 탑재한 모바일 기기 보급으로 위치 기반 정보는 모바일 생활의 필수 요소로 자리잡았다. 이제 사용자들은 더 나아가 개인별 성향에 따른 맞춤형 정보를 원하고 있다. 개인 맞춤형 추천을 위해서는 사용자의 행동을 이해하는 것이 필요한데 실생활에 많은 부분을 차지하고 있는 음식점 방문도 맞춤형 추천 서비스를 제공할 수 있다. 본 논문에서는 음식점 방문에 대한 비슷한 태도를 보인 사용자를 추출한 후 방문했던 장소를 비교하여 추천하는 사용자 행동 기반 다속성 태도 모델 기반의 장소 추천 모델을 연구한다. 다속성 태도점수를 산출하기 위해 피쉬바인(Fishbein) 방정식을 활용하고 피어슨 상관계수를 이용하여 사용자들이 방문했던 장소의 중 유사한 속성을 가진 장소를 추출했다. 그리고 그룹렌즈의 선호도 예측 알고리즘을 활용하여 추천 대상 장소를 선정하고 유클리디안 거리법으로 사용자에게 거리기반 장소를 추천하였다. 또한 실제 데이터를 이용한 실험을 통해 본 논문에서 제시한 시스템의 우수성도 입증하였다.

■ 중심어 : 사용자태도모델 ; 빅데이터분석 ; 사용자유사도 ; 장소추천

Abstract

For a place-recommendation model based on user's behavior and multi-attribute attitude in this thesis. We focus groups that show similar patterns of visiting restaurants and then compare one and the other. We make use of The Fishbein Equation, Pearson's Correlation Coefficient to calculate multi-attribute attitude scores. Furthermore, We also make use of Preference Prediction Algorithm and Distance based method named "Euclidean Distance" to provide accurate results. We can demonstrate how excellent this system is through several experiments carried out with actual data.

■ keywords : User Behavior-Based Multi-Attribute Attitude Models ; Big data Analytics ; User Similarity ; Place Recommendation System

I. 서론

모바일 기반의 추천시스템은 사용자의 위치와 장소의 거리를 계산하여 가까운 곳으로 추천하는 단순한 방식에서 위치정보에 대한 정확도 및 정보 보안 등에 대한 기술 중심적인 연구, 사용자 행위 관점에서의 연구, 새로운 커뮤니케이션의 수단으로 마케팅 관점에서의 연구 등이 이루어지고 있다.

최근에는 SNS(사회적연결망서비스)가 성장하면서 상호 관계가 있는 사용자들로부터의 장소 추천이 가능해지는 등 모바일 위치기반 장소 추천은 점점 진화하고 있다. 신뢰도를 바탕으로 하는 사회적 관계에서의 추천은 사람마다 각기 다른 취향이 있을 경우 추천받은 음식점을 방문했을 때 만족스럽지 못할 수 있다. 또한, 기존의 장소 추천 연구는 사용자가 관심영역을 직접 설정하는 등의 개인 사용자의 모델링으로 콘텐츠를 추천하는

* (주) 씨온 대표이사

** (주) 씨온 마케팅본부 본부장

*** 정회원, (주) 씨온 마케팅본부 기획팀

본 연구는 문화체육관광부 및 한국문화관광연구원의 2015년 관광 서비스 R&D 지원 사업의 연구결과로 수행되었음.

접수일자 : 2016년 06월 02일

게재확정일 : 2016년 06월 29일

수정일자 : 2016년 06월 22일

교신저자 : 최혜림, e-mail : haelim3790@naver.com

연구와 사회적 관계가 있는 사용자간의 체크인 유사도로 추천해주는 연구 등이 있다. 이 연구들은 사용자의 방문 기록만 가지고 장소를 추천했기 때문에 주관적인 평가가 포함되지 않아 만족스럽지 못할 수 있다.

따라서, 기존에 방문했던 음식점에 대한 사용자의 방문태도(행동) 점수를 반영하고 분석하여 그 태도가 비슷했던 사용자간의 유사도를 측정하여 장소를 추천하는 기술 개발 연구가 필요하다.

본 연구에서는 음식점 방문에 대한 비슷한 태도를 보인 다른 사용자를 추출한 후 방문했던 장소를 비교하여 추천하는 사용자 행동 기반 다속성 태도 모델 기반의 장소를 추천하는 연구를 해보려고 한다.

II. 본 론

1. 관련연구

1.1 사용자 다속성 태도모델

사용자가 방문한 장소에 대한 행동 즉 태도를 분석하기 위해 소비자 행동론에서 활용하고 있는 피쉬바인 다속성 태도모형을 활용하였다. 피쉬바인 모델에서는 소비자가 어떤 제품에 대해 갖게 되는 태도는 그 대상에 대해 가진 여러 정보 중 소비자가 지각한 몇 가지 현저한 신념, 그리고 브랜드 속성에 대한 중요성 평가에 의해 총체적 태도가 결정된다고 말하고 있다. 이 알고리즘은 소비자의 구매 경향에 대해 이해하는데 목적이 있으며, 구매 경향과 긴밀한 태도 점수를 구할 수 있어 소비자 태도 체계에 대한 이해를 얻을 수 있는 장점이 있다.

$$A_{u_1 p_1} = \sum_{i=1}^n b_i e_i \quad (1)$$

식 (1) 피쉬바인(Fishbein)의 다속성 태도모델 방정식

식 (1)은 장소 p_1 에 대한 사용자 u_1 의 다속성 태도 점수이며, n 은 관련 속성(ex. 친절도, 가격, 청결도 등)의 수를 의미한다. b_i 는 속성 i 에 대한 사용자 u_1 의 신념의 강도를, e_i 는 속성 i 에 대한 사용자 u_1 의 평가를 의미한다. 식 (1)을 통해 장소 p_1 을 방문한 사용자들은 속성에 따라 태도 점수를 계산할 수 있다.

1.2 협업적 필터링 기반 추천 프로세스

협업적 필터링은 과거에 아이템을 구매하거나 평가하는데 있어서 유사한 성향을 보였던 사용자가 경험한 아이템간의 비교로 본인이 이전에 구매하지 않은 아이템을 추천하는 것이다. 협업적 필터링 기반 알고리즘은 정량적으로 분석하기 어려운 아이템에 적용했을 때 좋은 성능을 나타내어 책을 추천하는 Amazon.com, 음반을 추천하는 CDnow.com등 상업적으로 성공한 전자상거래 사이트에서 적용하고 있다.

- 사용자 프로파일 구성

사용자의 이전의 체크인 데이터와 각 아이템에 대해서 평가한 점수를 이용하여 생성한다.

- 고객간 유사도 계산

고객이 각 장소에 평가한 점수를 이용해서 각 고객의 다른 고객과의 유사도를 구하는데 여기에는 피어슨 상관계수를 사용한다. 사용자 a 와 u 의 유사도는 다음과 같다.

$$Sim'_{(a,u)} = \frac{\sum_{i=1}^m (r_{ai} - \bar{r}_a)(r_{ui} - \bar{r}_u)}{(m-1)\sigma_a\sigma_u} \quad (2)$$

식 (2) 피어슨 상관계수 계산식

식 (2)는 사용자 a 와 u 가 공통으로 평가한 아이템을 바탕으로 계산한 두 사용자 간의 유사도를 나타내며, m 은 두 사용자가 공통으로 평가한 아이템의 수를 의미한다. 그리고 r_{ai} 는 예측하고자 하는 사용자 a 가 i 번째 아이템에 대해 수행한 평가 점수이며, r_{ui} 는 사용자 u 가 아이템 i 에 대해 수행한 평가 점수이다. 또한 \bar{r}_a 와 \bar{r}_u 는 각각 사용자 a 와 u 가 수행한 모든 아이템에 대한 평가점수의 평균이며, σ_a 와 σ_u 는 각각 사용자 a 와 u 가 수행한 평가 점수의 표준편차다.

- 선호도 예측

고객간 유사도를 이용하여 각 아이템 i 에 대한 사용자 a 의 예측치를 다른 사용자들의 평균으로부터 편차들에 대한 가중 평균을 이용하여 다음과 같이 구한다.

$$P(a,i) = \bar{r}_a + \frac{\sum_{u \in S(a)} Sim'(a,u)(r_{ui} - \bar{r}_u)}{\sum_{u \in S(a)} Sim'(a,u)} \quad (3)$$

식 (3) 사용자 선호 예측값 산출식

식 (3)의 $P(a,i)$ 는 사용자 a 가 경험하지 않은 아이템 i 에 대한 a 의 선호 예측값을 의미한다. \bar{r}_a 와 \bar{r}_u 는 각각 사용자 a 와 u 가 평가한 모든 아이템에 대한 평가 점수의 평균을 나타

내고, $r_{u,i}$ 는 이웃 사용자 u 의 아이템 i 에 대한 평가점수를 나타낸다.

1.3 사용자 태도모델과 협업적 필터링의 결합

본 논문에서는 사용자 행동 태도를 기반으로 협업적 필터링에서의 장소 추천 문제점을 보완한 장소추천 기법을 제안한다. 또한 거리계산 알고리즘을 추가하여 장소 추출에 최적화 하였다. 또한 선호점수와 거리 값에 가중치 계수를 두어 추출된 결과에 대해서 상황에 따라 추천 장소 기준을 조정 하게 하여 보다 유용하게 장소를 추천 할 수 있도록 하는 모델을 제시하였다.

2. 사용자 다속성 태도기반 장소추천시스템

본 장에서는 피쉬바인 방정식과 피어슨 상관계수, 협업적필터링 알고리즘, 유클리디안 알고리즘 등을 기반으로 하는 사용자 다속성 태도 기반의 장소추천 시스템의 주요 처리 과정에 대해서 설명한다. 그림 1은 사용자 다속성 태도 기반 장소 추천시스템의 처리과정을 보여준다.

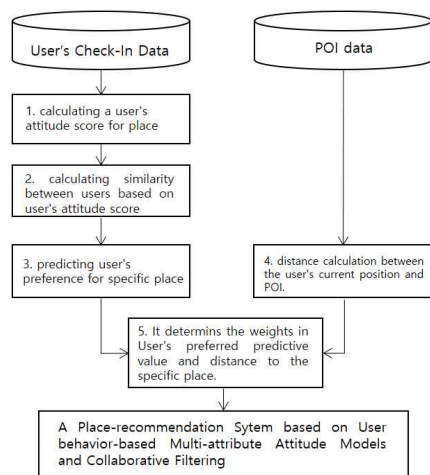


그림 1. 사용자 다속성 태도 기반 장소 추천시스템

그림 1은 사용자에게 장소 추천을 하기 위해 사용자가 온라인 상에서 체크인한 데이터를 분석하여 특정 장소에 대해 높은 선호도인 장소를 추출해 내는 과정이다.

본 논문에서는 특정 장소의 방문 태도 점수를 구하여 사용자 간의 유사도를 계산하고 유사도가 높은 사용자들이 방문했던 장소 간 유사도를 계산하여 선정된 선호도를 예측하겠으며 가까운 거리의 장소를 추천하도록 하겠다. 추천하는 식은 다음과 같다.

$$V_u = \chi \sum_{j=1}^m P_{uj} + \epsilon \log \frac{1}{\sqrt{\sum_{j=1}^m (CL_u - PL_{uj})^2}}, \quad \chi + \epsilon = 1 \quad (4)$$

식 (4) 사용자 선호 예측에 대한 거리 기반 장소 추천식

식 (4)의 V_u 는 u 사용자의 예측 선호도 기반 장소 추천 값으로 타 사용자와의 유사도와 장소 간 유사도로부터 계산된 u 사용자의 장소 선호 예측 값이 나온 장소 P 의 모든 장소들을 사용자의 현재 위치 CL_u 와 선호 예측된 장소의 위치 PL_{uj} 와 의 거리 계산 값을 역수로 했을 때 값이 크면 u 사용자의 선호가 예측되는 장소 중 가까운 장소로 추천한다. χ 와 ϵ 는 추천된 장소 결과의 상황에 따라 적용 가능한 가중치 계수다.

2.1 데이터 수집

사용자 다속성 태도를 기반으로 선호 장소를 예측하기 위해서는 방문 평가를 남긴 기록이 있어야 하며 방문 평가의 속성에는 장소 평가 점수와 태도 점수가 포함되어야 한다. 태도 점수는 음식점을 기준으로 친절도, 청결도, 맛, 가격 등을 정하도록 했다.

본 논문에서는 맛집 정보서비스인 S서비스의 방문평가 데이터를 활용하였다.

표 1. S사 사용자의 방문평가(체크인) 데이터

User's check-in data			
UserName	Jane.k	cScore	3
Pname	원조보쌈	kScore	5
LAT	37.5	tScore	4
LNG	127	pScore	4
ADDR	서울시 중구 장충동...	lScore	2
Category	족발/보쌈

체크인 데이터는 장소에 대한 이름, 좌표, 사용자가 체크인을 하면서 생성한 데이터인 닉네임, 장소평가점수, 태도점수로 구성되어 있다. 본 논문에서는 동일한 장소를 평가한 사용자들을 대상으로 선호 예측 장소 추천 모델을 적용시켜 보았다.

- 수집 조건: 2015년 10월 1일~2015년 12월 31일 기간 동안 S서비스에서 장소의 방문 평가 기록을 남긴 사용자 20명

2.2 사용자의 태도점수 산출

사용자의 태도점수를 산출하기 위해서는 추천하는 장소의 태도를 중요도에 따라 비율을 설정하고 사용자마다 태도에 대한 신념의 강도가 있어야 한다. 사용자의 신념의 강도는 평가 이전에 있어야 하며 S서비스에서는 장소를 최초로 평가 할 때 입력

받았다.

가. 사용자의 신념의 강도

S서비스를 기준으로 신념에 대한 강도를 입력한 사용자 데이터는 표 2와 같다.

표 2. 사용자 u_1 의 신념에 대한 강도

태도	강도 (중요도)	환산점수	평가점수	환산점수 x평가점수
cScore	1	0.4	3	1.2
kScore	2	0.3	5	1.5
tScore	3	0.15	3	0.45
pScore	4	0.1	1	0.1
lScore	5	0.05	1	0.05
태도 점수				3.3

나. 사용자의 태도 점수

사용자의 신념의 강도를 수치화하고 장소에 대해 태도 평가를 했으면 장소별 태도점수를 산출할 수 있다.

표 3. 사용자의 장소에 대한 태도 점수

구분	u1	u2	u3	u4	u5	...	u20
p1		4.7		2.7			
p2				1.8		3.05	
p3	3.3		4.2		2.55	4.12	
...	3.6		3.35				4.5
pj		3.45		3	3.05	3.55	3.35

표 3은 피쉬바인 방정식으로부터 사용자의 태도점수를 산출한 표다. 사용자들이 방문한 장소들에 만 점수가 표시되어 있으며 같은 장소를 평가한 사용자들 간의 유사도를 계산하여 비슷한 성향의 사용자를 예측할 수 있다.

2.3 태도 점수 기반 사용자간 유사도 계산

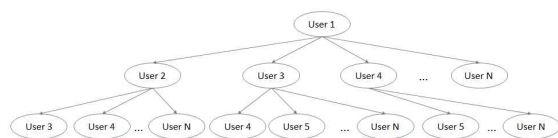


그림 2. 사용자 비교 모델

동일한 장소를 평가한 사용자들의 태도 점수로 피어슨상관 계수 식을 이용하여 비슷한 성향의 사용자들을 예측할 수 있다. 유사도 값 범위는 $-1 < x < 1$ 이며, -1 에 가까우면 유사도가 낮다고 볼 수 있다. 표 4는 사용자간 유사도를 산출한 결과다.

표 4. 사용자간 유사도 산출값

사용자		유사도
User1	User2	0.641866715
	User3	0.692397862

	User20	0.816194389
User2	User3	-0.390463908
	User4	-0.553396111

	User20	-0.034241188
.....		
User19	User20	-0.07051282

2.4 선정된 장소의 사용자 선호 예측

예측치는 사용자들의 선호도를 가중 평균하여 평가 점수로 예측하며, j 번째 장소에 대한 사용자 u_1 의 예측치는 다음과 같은 식으로 나타낼 수 있다.

$$P(u_1, p_j) = \bar{r}_{u_1} + \frac{\sum_{u \in S(u_1)} Sim'(u_1, u_{20})(r_{u_{20}, p_j} - \bar{r}_{u_{20}})}{\sum_{u \in S(u_1)} Sim'(u_1, u_{20})} \quad (5)$$

식 (5)의 $P(u_1, p_j)$ 는 사용자 u_1 이 경험하지 않은 장소 p_j 에 대한 선호 예측값을 의미한다. \bar{r}_{u_1} 과 $\bar{r}_{u_{20}}$ 은 각각 사용자 u_1 과 u_{20} 이 평가한 모든 아이템에 대한 평가 점수의 평균을 나타내고, r_{u_{20}, p_j} 는 이웃 사용자 u_{20} 의 p_j 에 대한 평가점수를 나타낸다.

모든 사용자를 목표 사용자의 이웃으로 지정 하는 것이 아니라, 유사도 임계값 (본 연구에서는 0.8로 정의한다.)을 지정하여 그 이상의 값을 가진 사용자만을 목표 사용자의 이웃으로 지정한다. 근접 이웃 사용자의 평가 값, 유사도만을 사용하여 목표 사용자의 평가되지 않은 장소에 대한 선호 예측값을 계산한다.

다음 표는 사용자 u_1 과 유사도가 가장 높은 사용자 u_{20} 이 평가한 장소에 대한 u_1 의 선호 예측값이다..

표 5. 사용자 u_1 과 유사도가 가장 높은 u_{20} 이 방문한 장소에 대한 u_1 의 선호 예측값

User	u_{20} 이 방문한 장소	Prediction
User u1	Place11	2.313
	Place15	4.153
	Place17	3.973
	Place24	4.153
	Place28	4.483
	Place32	3.253

	Place36	3.983
	Place38	3.383

2.5 사용자의 현재 위치와의 거리 값을 계산한 추천 대상 장소의 값

선호 예측 값이 높은 순으로 추천 대상인 장소를 정렬하고 사용자 u 의 현재 위치 좌표와 계산하여 가까운 장소로 최종 추천한다.

사용자의 현재 위치와 추천 대상인 장소의 거리의 차이를 계산한 값은 소수점으로 나옴으로 역수로 하여 값이 큰 장소를 사용자 u 의 현재 위치와 가까운 것으로 본다.

예측 점수와 거리 점수를 더하면 추천 대상 장소의 값을 구할 수 있다.

표 6. 사용자 선호 예측값과 거리에 따른 추천 장소

Pname	Prediction	Distance Score	추천대상 장소의 값	
			(선호점수 : 거리 = 0.8 : 0.2)	추천 순위
Place28	4.483	3.254896	4.2373792	1
Place15	4.153	2.549756	3.8323512	4
Place24	4.153	3.648751	4.0521502	2
Place36	3.983	1.248756	3.4361512	5
Place17	3.973	3.669874	3.9123748	3
Place38	3.383	2.459785	3.198357	7
Place32	3.253	3.045787	3.2115574	6
Place11	2.313	2.165775	2.283555	8

값이 높은 장소가 사용자의 태도를 기반으로 추천한 장소라고 볼 수 있다.

3. 구현 및 실험

3.1 시스템 개발

본 논문에서는 S서비스의 체크인 데이터를 활용하여 스마트폰 기반의 안드로이드 OS에서 구현해보았다. 그림 3은 본 논문에서 구현한 실험 화면을 보여 준다.

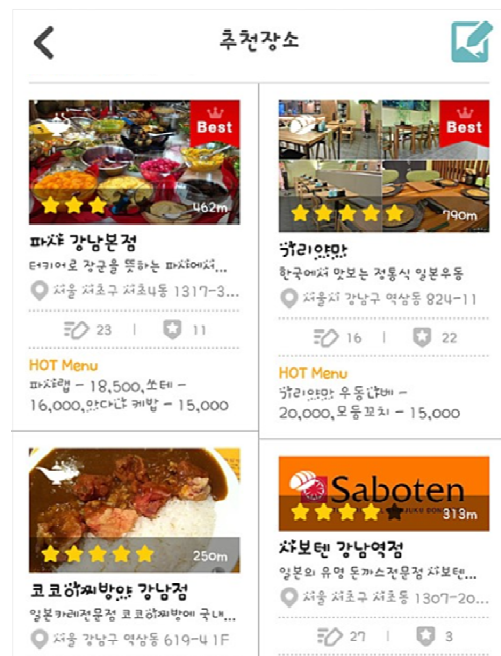


그림 3. 사용자 행동 기반 장소 추천 시스템

3.2 만족도 비교

본 논문에서 제안하는 방법과 기존 논문[4]의 추천 방법을 비교해보았다. 비교 실험을 위해 음식점 추천 모바일 어플리케이션을 이용하는 실험자 20명을 대상으로 10개 장소에 대해 만족도 평가를 진행하였다.

그래프에서 previous method로 표기된 선은 기존 논문[4]에서 사용한 방법에 따른 장소별 사용자 만족도 평가 결과를 나타내며, proposed method로 표기된 선은 식(4)에 기술된 방법에 따른 장소별 사용자 만족도 평가 결과를 나타낸다.

그림 4를 보면, LBSNS의 체크인 데이터 중 방문했던 음식점 카테고리를 기반으로 협업적 필터링과 피어슨 상관계수로 사용자간 유사도를 측정하여 장소를 추천하는 방법[4]보다 만족도가 40% 향상되는 것을 알 수 있다. 따라서 본 연구에서 제안하는 모델로 사용자의 맞춤형 장소를 추천할 수 있다.



그림 4. 제안하는 방법과 이전 방법 비교

III. 결 론

본 논문에서는 위치기반서비스와 빅데이터 기술이 결합되어 발전되고 있는 사용자 행동 기반 추천 시스템을 연구하였다. 본 논문에서 개발한 사용자 행동 기반 다속성 태도 모델기반의 유사도 측정 연구는 사용자 신념에 따른 태도 점수, 사용자간 유사도, 협업적 필터링 선호예측, 유클라디안 알고리즘 등을 기반으로 개인화에 따른 맞춤형 장소를 추천할 수 있다.

본 논문에서 제시한 방법을 이용하여 장소를 추천할 경우, 장소에 대한 방문 태도를 분석하고 본인과 방문 태도가 유사한 사람들이 방문한 장소를 추천 받기 때문에 개인화된 맞춤형 장소 추천이 가능해져 보다 높은 만족도로 장소를 방문할 수 있게 되었다. 또한 실제 사용자의 만족도 평가를 통해 기존 연구 방법보다 우수하다는 것을 입증하였다. 하지만 점수만으로는 개인의 성향을 파악하기에 부족할 수 있기 때문에 보다 정교한 추천을 위해서는 음식점 방문 평가 후기를 기반으로 한 추천 시스템 연구가 필요할 것이다.

References

- [1] 김유정, 김광지, 박기용, “피쉬바인 모델을 이용한 커피전문점의 소비자 태도에 관한 연구”, *한국조리학회지*, 제17권, 제5호, pp33-38, 2011
- [2] 남일우, “협업적 필터링을 이용한 쇼핑 추천 시스템”, *인하대학교 대학원 석사학위논문*, pp6-15, 2014
- [3] 박지선, 김택헌, 류영석, 양성봉, “추천 시스템을 위한 2-way 협동적 필터링 방법을 이용한 예측 알고리즘”, *정보과학회논문지 소프트웨어 및 응용*, 제29권, 제9·10호, 2002년 pp669-675
- [4] 신흥철, “쿼드트리를 적용한 위치 기반 협업 필터링 추천 시스템”, *연세대학교 공학대학원 석사학위논문*, pp6-12, 2012
- [5] 오세창, 최민, “항목 간 선호도 차이를 이용한 영화 추천 방법”, *한국정보통신학회논문지*, pp2603-2606, 2013
- [6] 윤성태, “패밀리레스토랑의 청결라벨유무 및 청결라벨인지도에 따른 소비자태도의 비교”, *세종대학교 석사학위논문*, 2012
- [8] 윤소영, 윤성태, “사용자 정보 가중치를 이용한 추천 기법”, *한국해양정보통신학회논문지*, 제15권, 제4호, pp879-887, 2011
- [9] 이락규, 피준일, 박준호, 복경수, 유재수, “모바일 환경에서 콘텐츠 추천 시스템 설계 및 구현”, *한국콘텐츠학회논문지*, pp45-47, 2011
- [10] 정광재, 송용택, “모바일 인터넷 이용자 및 이용패

턴 특성과 인터넷 중립성”, *정보통신정책연구원*, pp48-51, 2014

- [11] 정구임, “LBSNS 기반 장소 추천 방법 연구”, *건국대학교 정보통신대학원 석사학위논문*, pp23-38, 2012
- [12] TIAN LIANHUA, “소셜 정보 기반의 맛집 추천 시스템”, *이화여자대학교 대학원 석사학위논문*, pp24-43, 2013

저 자 소 개



안 병 익 (Byung-IK Ahn)

1993년 동국대학교 컴퓨터공학과 공학석사
2007년 연세대학교 컴퓨터과학과 공학박사
2010년 5월 ~ 현재 (주) 씨온 대표이사

<주관심분야 : 컴퓨터사이언스, 모바일컴퓨팅, 소셜컴퓨팅>



정 구 임 (Ku-Imm Jung)

2013년 건국대학원 정보통신학과 공학석사
2010년 7월 ~ 현재 : (주)씨온 마케팅본부 본부장

<주관심분야 : LBS, LBSNS 플랫폼, 빅데이터 분석 및 추천>



최 혜 림 (Hae-Lim Choi)

2013년 건국대학교 인터넷미디어공학부 멀티미디어공학과 학사
2014년 1월 ~ 현재 : (주)씨온 마케팅본부 기획팀 사원

<주관심분야 : 빅데이터 분석 및 추천, 데이터베이스, 모바일컴퓨팅>